Disaster Dynamics: Understanding the Role of Quantity in Organizational Collapse

Jenny W. Rudolph
Boston College
Nelson P. Repenning
Massachusetts Institute of Technology

This article examines the role that the quantity of non-novel events plays in precipitating disaster through the development of a formal (mathematical) system-dynamics model. Building on existing case studies of disaster, we develop a general theory of how an organizational system responds to an on-going stream of non-novel interruptions to existing plans and procedures. We show how an overaccumulation of interruptions can shift an organizational system from a resilient, self-regulating regime, which offsets the effects of this accumulation, to a fragile, self-escalating regime that amplifies them. We offer a new characterization of the conditions under which organizations may be prone to major disasters caused by an accumulation of minor interruptions. Our analysis provides both theoretical insights into the causes of organizational crises and practical suggestions for those charged with preventing them.●

Major disasters have long interested organization theorists (Perrow, 1984; Shrivastava, 1987; Weick, 1993b; Vaughan, 1996), and their causes continue to be an active area of inquiry. Accidents like the nuclear catastrophe at Chernobyl or Union Carbide's gas leak at Bhopal are major social events responsible for immeasurable human suffering and environmental damage. There are few more compelling opportunities for organization theory specifically, and the social sciences in general, to prevent suffering and contribute to humanity. Moreover, major disasters provide a unique opportunity to study organizational processes in situations that are far from equilibrium. Just as the designers of bridges and airplanes test their systems under extreme conditions that are rarely, if ever, experienced during actual use, major catastrophes provide a similar opportunity to learn more about the vulnerability and resilience of human and social systems.

The literature on disaster and its flip side, safety, includes in-depth case studies (e.g., Shrivastava, 1987; Weick, 1993b; Vaughan, 1996), studies of learning from accidents and error (e.g., Cook and Woods, 1994; Carroll, 1995), theories of high-hazard or accident-prone organizations (Turner, 1976; Sagan, 1993; Perrow, 1994), theories of high-reliability organizations (Roberts, 1990; Schulman, 1993; Weick, Sutcliffe, and Obstfeld, 1999), and theories of how to manage accident and error (e.g., Reason, 1997). A significant insight emerging from this literature is that major disasters often do not have proportionately large causes. Theorists increasingly recognize that small events can link together in unexpected ways to create disproportionate and disastrous effects (Weick, 1993a; Perrow, 1994; Vaughan, 1996; Reason, 1997). Perrow (1984) suggested that as production technologies become increasingly sophisticated and interconnected with other systems, the likelihood of chain reactions, in which one problem reverberates through the system and triggers a cascade of malfunctions and breakdowns, greatly increases the chance that minor, everyday events will lead to major disasters, or what he called "normal accidents."

An important implication of Perrow's normal-accident theory is that complex, tightly coupled systems often produce cues that are either invisible or defy existing categories. The chal-

0001-8392/02/4701-0001/$3.00.

●

For their helpful comments, thanks to Debra Ancona, Paul Carlile, John Carroll, Erica Foldy, Paulo Goncalves, Lee Fleming, Danna Greenberg, Candace Jones, Mark Kleiman, Tammy MacLean, Brad Morrison, James Reason, Peter Repenning, John Sterman, Steve Taylor, Bill Torbert, Diane Vaughan, Karl Weick, Joanne Yates, seminar participants at Harvard, MIT, and Boston College, three anonymous reviewers, the associate editor, Joe Porac, and the managing editor, Linda Johanson.

lenges of managing such events are acknowledged both by proponents of this high-hazard view and those in the countervailing high-reliability school (cf. Weick, Sutcliffe, and Obstfeld, 1999). While the two schools disagree on the ability of organizations to handle this challenge, they both emphasize the central role of novel events in precipitating crisis. In situations ranging from anomalous O-ring data leading to the space shuttle disaster (Vaughan, 1996) to a wandering bear almost precipitating nuclear war (Sagan, 1993), novel events that challenge conventional categorization or response often play an important role in major disasters.

Both social psychological and sociological analyses of disasters often focus on the processes through which novel events are sensed and resolved. For example, organizations fail to perceive novelty when it is embedded or obscured by complex technology (Perrow, 1994), when the complexity of the external environment outstrips the organization's ability to sense it (Weick, 1993a), or when the novelty is so extreme that it cannot be accommodated in the existing worldview (Weick, 1993b). Similarly, even when novelty is perceived, it is often suppressed, particularly when acknowledging it undermines existing organizational goals or norms (Turner, 1978; Shrivastava et al., 1988; Vaughan, 1996). Policy prescriptions emerging from these literatures include widening attention and conceptual categories; simplifying complexity, yet doubting those simplifications; and being willing to reframe perceptions on the fly (Carroll, 1995; Weick, Sutcliffe, and Obstfeld, 1999).

Highlighting novelty as a basis for disaster represents an important intellectual milestone in understanding organizations and their vulnerabilities. Like all powerful analytical approaches, however, it has led to blind spots (cf. Weick, 1979). A careful reading of existing work suggests that in many cases both the novelty and the quantity of interruptions to established routines and expectations play a significant role in precipitating disaster. For example, the anomalous O-ring performance (novelty) that ultimately caused the *Challenger* disaster was just one of many outstanding risk-tolerance issues (quantity) that had to be resolved within an allotted time (Vaughan, 1996). Similarly, the bear mistaken for a Soviet saboteur (novelty) happened to enter the compound at the height of the Cuban missile crisis, while U.S. soldiers were carrying out the numerous tasks needed to mobilize bombers (quantity) in anticipation of nuclear war (Sagan, 1993). Yet, despite its frequent appearance in situations that ultimately end in disaster, the role of quantity has received relatively little attention.

The links between non-novel disruptions and organizational collapse may have been given short shrift because the dynamics seem obvious: the more there are, the worse things get. While it is tempting to invoke such proportional logic to explain the role of overload, there is reason to suspect that the dynamics are more complex. While Perrow characterized physical systems as either loosely or tightly coupled, implying that vulnerability to small breakdowns is a fixed feature of a given system, Weick (1993a: 189) and others have suggested that the ability of human systems to accommodate interruptions and breakdowns without descending into crisis is both contextually and temporally dependent: "when you take people and their limitations into account, susceptibility to [disasters caused by minor events] can change within a relatively short time." Similarly, working from a sociological perspective, Turner (1978: 89) argued that the accumulation of unnoticed events during an "incubation" period can cascade into disaster as the result of a precipitating event.

While these observations suggest the importance of contextual and temporal factors in determining a system's vulnerability to crisis, they have overlooked the role that the quantity of non-novel disruptions plays in this process. Little theory currently exists to understand the dynamic interplay among the quantity of small events, the state of the surrounding environment, and the organization's capacity for responding to it. Without a strong theoretical characterization of these dynamics, it is difficult to determine whether the strategies so helpful in offsetting the potentially damaging effect of novel interruptions are similarly effective when quantity is also an issue. This study thus examines the interconnections between the quantity of small, non-novel events and organizational crises.

Building on in-depth analyses of a number of existing case studies and relevant psychological theory, we propose and analyze a dynamic model of an organizational system facing an ongoing stream of interruptions. Our model synthesizes causal processes found in existing analyses of disaster and experimental studies of human performance under stress to create a laboratory for studying and theorizing about the connections between quantity and disaster. Through simulation, we use our model to induce an internally consistent theory of the relationship between the quantity of interruptions and organizational performance. Our analysis provides a new characterization of how organizational systems respond to an ongoing stream of non-novel, survival-threatening interruptions.

METHODS

We explored the connections between small events and the collapse of an organizational system by developing a mathematical model. Unlike many formal models in the social science literature, ours was not deduced from general principles but, using the methods of grounded theory, was induced from theories and data from a range of domains. While commonly used to build theory from raw data using qualitative analysis, the grounded theory approach is not limited to this activity. Strauss and Corbin (1994) advocated the development of formal (or general) theories grounded in previously generated domain-specific (what they call substantive) analyses. They reminded the reader that Glaser and Strauss (1967) not only urged the use of grounded theory in conjunction with quantitative (not just qualitative) analysis but also recommended its use to generate theory from theory.

The purpose of our effort was to move toward a general explanation of how small events can create crises. We chose formal modeling as a tool for theory development because, while there is a rich array of narratives and case-specific theories of disaster, there have been fewer efforts to develop theory that abstracts from different domains (notable exceptions include Vaughan, 1999; Weick, Sutcliffe, and Obstfeld, 1999). Inducing a formal mathematical model facilitates the identification of structures common to the different narratives and enforces the internal consistency of the emerging theory. The translation of a narrative theory to a mathematical model results in some loss of richness and the ability to evoke nuances in organizational experience. The corresponding benefit is an internally and dynamically consistent theory whose core structure is explicitly represented.

The genesis of our theory was Weick's (1993a) description and analysis of the 1977 disaster at the airport in Tenerife. Using the logic of grounded theory building (Strauss and Corbin, 1994), we treated Weick's theoretical analysis as our initial source data and began our effort by translating his text-based constructs and theoretical relationships into the system-dynamics language of stocks, flows, and feedback loops (e.g., Forrester, 1961; Sterman, 2000). Following this mapping, we compared our diagrams with constructs and relationships identified in other case studies of disaster and empirical research on stress and interruptions in an iterative process of model elaboration and revision. During this step, the Yerkes-Dodson Law, which posits an inverted U-shaped relationship between stress and performance, emerged as central to understanding the role of quantity in organizational crises. We translated this emerging set of relationships into a formal mathematical model and then used computer simulation to analyze it. Finally, we returned to the literatures on stress, interruptions, and disaster, noting both analogies and differences. The result is a theory explicitly linking stress and performance that, while tightly grounded in previous work, reaches new insights concerning the role of interruptions in precipitating organizational collapse.[1]

QUANTITY AND CRISIS: TWO CASE STUDIES

To highlight the role that quantity plays in precipitating organizational collapse, we present two short case studies. The first is Weick's (1993a) vivid depiction of a series of small interruptions, none of which was particularly novel in and of itself, that combined to produce the Tenerife air disaster. On March 27, 1977, two Boeing 747s, one from KLM and one from Pan Am, were diverted to Tenerife because the Las Palmas airport, where they had been scheduled to land, was closed due to a terrorist bomb attack. Weick's analysis highlights how the diversion resulted in a myriad of small interruptions to existing plans and normal procedures: diverting the plane to Tenerife interrupted the plan to get back to Amsterdam within the KLM crew's strict duty time constraints; a cloud drifting 3000 feet down the runway interrupted the lower-order plan to leave the airfield; narrow runways (not designed for 747s) interrupted normal maneuvering protocols; and non-standard and garbled transmissions from the control tower interrupted usual preflight communications.

Invoking George Mandler's interruption theory of stress, Weick (1993a: 180) suggested that each of these interrup-

---

[1] A technical appendix to the paper containing complete model documentation, simulation instructions and a running version of the model is available at <web.mit.edu/nelson/www/Disasters.5.0.html>. The model is written and analyzed using the Vensim software, available from Ventanna Systems; see <www.Ventanna.com>.

tions increased the level of autonomic arousal in the KLM crew, absorbing information processing capacity, decreasing cognitive efficiency, and reducing the number of cues they were able to notice and process. As the situation progressed and the number of interruptions accumulated, the crew's ability to manage the increasingly complex system they were facing declined. The KLM crew communicated less and less clearly and developed a narrow and incomplete view of their situation, until, in direct violation of standard procedure, the KLM captain cleared himself for take-off. Then, to outrun a cloud rolling up the runway toward him, he began accelerating for take-off. Unfortunately, the approaching cloud concealed the Pan Am aircraft, which had missed its parking turn-off due to the low visibility. The resulting collision killed all of the 583 people on both planes, one of the worst accidents in aviation history.

The second case details the downing of a passenger jet by the USS *Vincennes* (see U.S. Senate Committee on Armed Services, 1988; Barry and Charles, 1992; Roberts and Dotterway, 1995; Klein, 1998; Collyer and Malecki, 1998). On July 3, 1988, the USS *Vincennes,* an AEGIS-system guided-missile cruiser designed to protect aircraft-carrier battle groups, mistakenly shot down a commercial airliner over the Persian Gulf, killing all 290 people on board. The shoot down was a tragedy for the people on Iran Air 655 and their families and a social disaster for the navy, unleashing a torrent of critical media attention and calling into question the navy's staff training, equipment design, and rules of engagement.

The role of cognitive biases, ambiguities in the engagement, the personality of the captain, and display design have all been emphasized in previous analyses of the event. Though not closely examined in previous analyses, the moments preceding the shoot down were punctuated by a continuing stream of new demands and breakdowns in ongoing processes that made it virtually impossible to carry out any one task without it being interrupted by another. The approaching unknown aircraft (Iran Air 655) interrupted the team's efforts to manage ongoing gun battles, and, given the plane's trajectory, the AEGIS combat-information-center team had just seven minutes to determine a course of action. High noise levels and bursts of information about the ongoing gunboat battles interrupted efforts to address the incoming aircraft. A jammed forward gun interrupted the ongoing battle maneuvers. Headphone communications reached crew members over several channels, with left and right ears sometimes receiving different messages, and the messages were periodically interrupted when everyone changed communication channels. This barrage of small interruptions raised stress levels and degraded the crew's cognitive and emotional capabilities. One careful analysis questioned "whether even the best-trained crew could handle, under stress, the torrent of data that AEGIS [information systems] would pour on them" (Barry and Charles, 1992: 33).

Building on these and similar examples, we developed a model to analyze the role that the quantity of interruptions plays in determining the resilience of organizational systems. We did not intend to capture the nuances that distinguish individual, group, and organizational responses. Instead, following Staw, Sandelands, and Dutton's (1981: 516) suggestion that such fine-grained analyses can "restrict our ability to see general patterns across social entities," we offer a single model that explicates a set of dynamics present at multiple levels of analysis. Following the presentation of our main results, we discuss how the processes we analyzed play out at different levels of analysis.

MODEL STRUCTURE

Interruptions

The central construct in our model is the interruption. We focus our theorizing on interruptions for two reasons. First, in the literature on disaster, interruptions to ongoing activities, plans, cognitive structures, or emotional gestalts appear repeatedly as a "generic accompaniment" of crisis (e.g., Staw, Sandelands, and Dutton, 1981; Perrow, 1984; Weick, 1993a: 182; Vaughan, 1996). Second, as our model emerged, Mandler's (1982) interruption theory of stress strongly shaped our view of how crises evolve. We stayed close to Mandler's theory (1982: 92) and adapted his definition: an interruption is any unanticipated event, external to the individual, that temporarily or permanently prevents completion of some organized action, thought sequence, or plan.

Our model is based on a number of key assumptions. We began developing it by assuming that the organization faces a continual (and potentially varying) stream of non-novel interruptions. In contrast to existing analyses, which focus on novelty, to study the dynamics generated by quantity, we restricted our attention to interruptions for which the organization has an appropriate response within its existing repertoire, our definition of a non-novel interruption. We did not model how the system might deal with novel interruptions (those for which responses in the existing repertoire are inappropriate). We revisit this distinction in the discussion, using it as a springboard to broader theorizing on the appropriate response to interruptions that require a mix of existing responses and newly invented ones.

Our focus on non-novel interruptions does not imply that all such events are created equal or that they can be resolved without significant cognitive effort. Resolving a non-novel interruption often requires both a shift to an active mode of cognition (Louis and Sutton, 1991) and the execution of at least three processes: attention processes to determine which interruptions are considered; activation processes to trigger the knowledge necessary in the given setting; and strategic processes to determine which goals are given priority and the resources that are allocated to them (Cook and Woods, 1994). In other words, members of the organization often have to stop and think before an interruption can be successfully resolved, and the difficulty of identifying and carrying out the appropriate response can vary widely with the interruption. For example, suppose that while a person is teaching or giving a talk, the stream of interruptions constitutes math problems of varying degrees of difficulty. A problem requiring addition or subtraction is easily handled, thus representing a fairly low-cost interruption. More complex problems, such as extrapolating exponential growth rates, while still within the capability of many people, are more time consuming and cognitively taxing (e.g., one must realize the problem can't be solved exactly without a calculator, identify an appropriate heuristic, and execute the heuristic).

To capture the idea that the responses to some interruptions are more ready at hand and thus require less time to identify and execute than others, we interpreted the interruption "units" in our model not as the raw number of interruptions but as the number of mental steps that resolving the associated interruption requires. For example, handling a new airplane in the approach path probably requires few steps for a seasoned controller, representing an interruption of just a few units. In contrast, responding to an airplane experiencing severe mechanical problems requires many more steps to identify and execute the appropriate response and, thus, is captured by many interruption units.

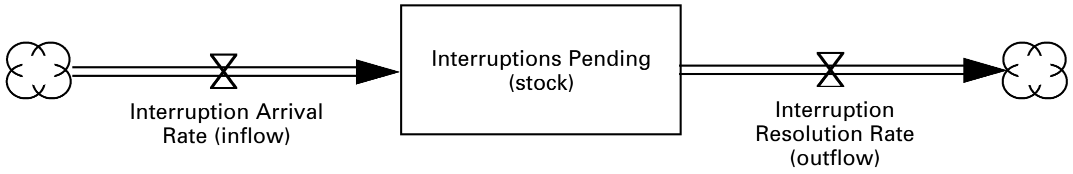
We also assumed that resolving pending interruptions is required for the individual's or organization's survival. Survival is broadly defined to include both physical functioning (e.g., the plane keeps flying) and the maintenance of social status and societal role (e.g., questions during a job talk must be answered to maintain one's status as a viable job candidate). Interruptions such as e-mail messages that disrupt research in progress, while annoying, do not immediately threaten either physical or social survival, and therefore fall outside the scope of our analysis. Finally, we did not model the primary task being interrupted but assumed that organizational performance is strictly a function of the ability to resolve interruptions.

How Interruptions Accumulate and Dissipate: Stock and Flow Structure

The process through which interruptions arrive, accumulate, and dissipate is shown in figure 1 in the form of a stock and flow diagram (Sterman, 2000), in which flow variables are signified in the diagram as "pipes" with "valves." The stream of incoming interruptions is represented as a flow variable, labeled the *interruption arrival rate.* Interruptions are not processed instantaneously but, instead, accumulate in the stock of *interruptions pending* (stocks are represented as rectangles). The stock, or level, represents the accumulation of interruptions that have occurred but have yet to be resolved. Formally,

$$\text{Interruptions pending}(t) = \int_t [\text{Interruption arrival rate}(s) - \text{Net interruption resolution rate }(s)]ds + \text{Interruptions pending}(t_0)$$

Figure 1. The basic stock and flow structure of interruptions in organizations.



The stock of interruptions pending is then reduced by the outflow, *net interruption resolution rate.*

That unresolved interruptions can accumulate follows directly from three aspects of our conceptualization: (1) resolving interruptions requires conscious cognitive effort; (2) the ability to attend to and execute tasks is finite; and therefore (3) the rate at which interruptions arrive may exceed the rate at which they are resolved. Further, because we assume that resolving such interruptions is necessary for ongoing survival, unresolved interruptions cannot be safely ignored. Instead, as work on cognitive task analysis highlights, participants engage in "mental bookkeeping," attempting to track unresolved interruptions until they are successfully handled (Cook and Woods, 1994). Examples of accumulated unresolved interruptions include requests from the control tower that have yet to be executed and questions in a job talk that have yet to be answered.

To define the net interruption resolution rate, we distinguished between interruptions and errors. People make mistakes, particularly when they are under pressure. Interruptions in our analysis were treated as exogenous inputs arising from outside the system that must handle them. In contrast, errors are endogenous—as the number of interruptions increases, errors become more likely—and can increase the probability of disaster in a number of ways. For example, errors made in complex, interconnected production systems have a tendency to reverberate, dramatically increasing the novelty of the task facing those attempting to manage it (Perrow, 1984). Given our focus on quantity, we do not capture these effects of errors in our model, but errors also affect the quantity of interruptions faced by the organization. When an error is made, even if it does not reverberate through the underlying technology, it often requires additional attention that would have been unnecessary had the original task been executed correctly. Errors often create additional interruptions. For example, a misinterpreted tower transmission leads to an incorrect response and causes the tower to repeat the communication; the interruption remains pending until correctly handled. Similarly, a poorly answered question in a job talk often leads to follow-up questions. To capture the effect of errors on the quantity of outstanding interruptions, we defined the outflow from *interruptions pending* as the rate at which interruptions are successfully resolved, labeling it the net, rather than the gross, resolution rate. Interruptions handled incorrectly simply remain in the stock of interruptions pending, creating potentially counterproductive stress, until they are successfully resolved. In using such a formulation, we assumed that errors are discovered immediately—an incorrectly handled interruption, rather than leaving the stock of interruptions pending and returning at some later moment when it is discovered, simply remains in the stock of interruptions pending.

While simple, the stock and flow structure of our model captures an important feature of organizational life largely ignored in other analyses: unresolved interruptions do not disappear. In lab studies of stress and performance (e.g., Yerkes and Dodson, 1908; Miller, 1978), stress is an independent variable manipulated by the experimenters and, within a given treatment, is constant. In such experiments, there is no relationship between the level of stress facing subjects and their past performance. While exogenously manipulating the level of stress facilitates accurate statistical estimation of the relationship between stress and performance, there is no such decoupling outside the laboratory. In many, if not all, organizational settings (and in our model), unresolved interruptions do not disappear but, instead, remain outstanding, continually producing stress until they are either forgotten (which takes time) or resolved. Thus, the level of stress facing an organization at any given moment is in large part determined by its past performance. In many real-world situations (as opposed to laboratory experiments), stress and performance, rather than being mono-causally related, are part of a feedback system in which current performance is affected by the number of unresolved interruptions, and the number of unresolved interruptions is determined, in part, by past performance. As we illustrate below, this feedback has important implications for understanding how the accumulation of interruptions determines organizational performance.

Linking Interruptions and Performance via the Yerkes-Dodson Law

To capture the linkage between the stress created by a large stock of unresolved interruptions and performance, we draw on the Yerkes-Dodson law, which posits an inverted U-shaped relationship between stress and performance on moderate to difficult tasks (Miller, 1978; Mandler, 1984; Fisher, 1986).[2] The Yerkes-Dodson law has a long and somewhat controversial history in psychological research. While originally derived from applying electrical shocks of different intensities to mice running through a maze, the curve has since been invoked to describe the effects of anxiety, arousal, drive, motivation, activation, and reward and punishment on performance, problem solving, coping, and memory in animals and humans (Teigen, 1994). Critics (e.g., Teigen, 1994; Bauemler, 1995) have contested both the replicability of these findings in domains beyond the original experiments and suggest that its wide range of application indicates a problematic level of vagueness in the underlying psychological constructs.

Despite these challenges, a close review of this literature suggests that, because we focused on situations in which the rapid resolution of interruptions is required for physical or social survival, the Yerkes-Dodson curve is an appropriate depiction of the relationship between stress and performance in the situations we studied. Empirical studies that fail to support the curvilinear hypothesis often look at chronic stressors (like role ambiguity, role overload) over long time periods (weeks or months), rather than situations in which interruptions must be resolved in minutes, or at most, a few hours (see Sullivan and Bhagat, 1992, for a recent survey). Studies that focus on tasks executed under significant time pressure, situations in which autonomic arousal is likely to be a significant factor, tend to support the curvilinear hypothesis (e.g., Coles, 1974). Moreover, we studied the effect of accumulating interruptions (stress) on the ability to resolve interruptions

---

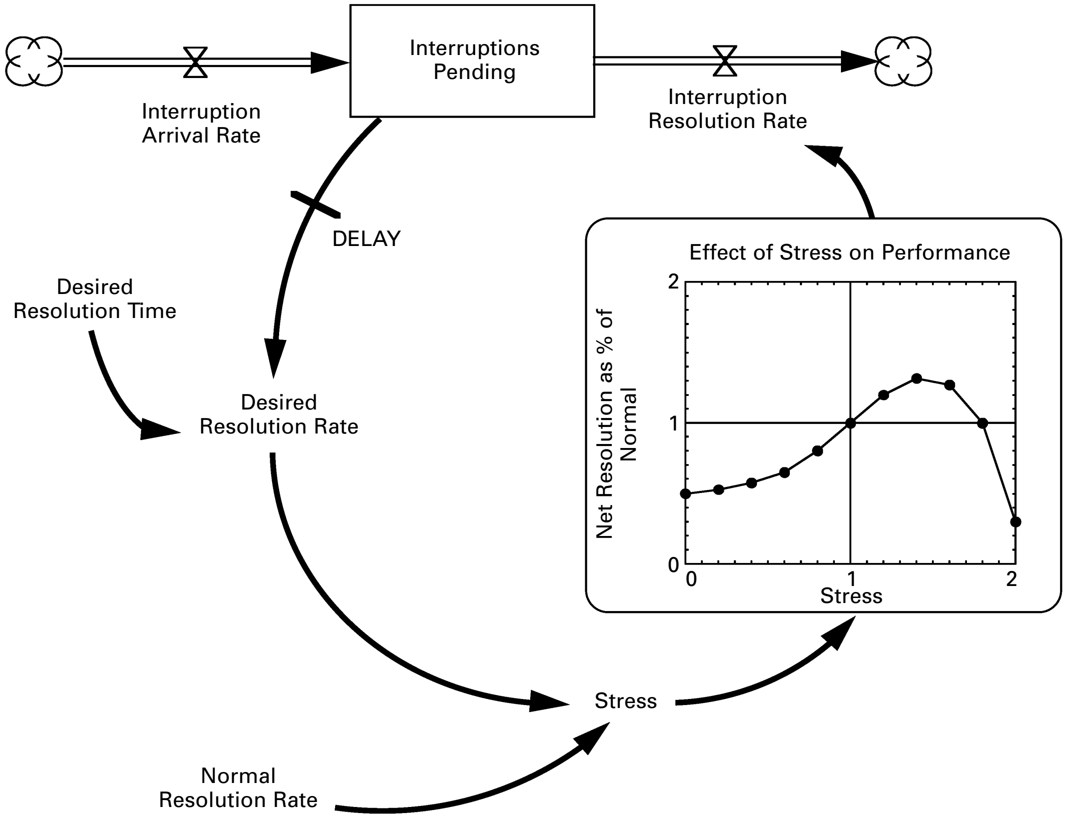
[2] Research on cognition and affect (e.g., Revelle and Loftus, 1990) indicates that affect, like arousal, has an impact on the encoding, retrieval, and utilization of information. In our model, we hold affect constant and focus only on changes in arousal.

successfully. Errors reduce performance in the contexts we studied, since they increase the number of outstanding interruptions. Once performance is defined as the rate at which information is processed correctly, the curvilinear relationship between the number of tasks pending and performance is even more robust (Miller, 1978). A large number of studies on the ability of individuals to process information (the closest experimental analog to the situation we studied) have found an inverted U-shaped relationship between the rate of information inputs and the ability of individuals to produce correct responses (see Miller, 1978).

Cognitive theory further supports the notion that a growing stock of pending interruptions impairs the execution of all the cognitive processes necessary for resolving them. A burgeoning stock of pending interruptions increases and complicates the tasks of allocating and controlling attention. Mental bookkeeping becomes increasingly difficult as the number of interruptions being tracked grows, decreasing awareness of the bigger situational picture. Losing situation awareness can lead to an erroneous and increasingly tight fixation on a particular framing of the situation, which, in turn, limits the perception of disconfirming cues (De Keyser and Woods, 1990; Rudolph, 2002). Accumulating interruptions also decrease the time available to process available stimuli and extract the patterns necessary to trigger inert knowledge (Gentner and Stevens, 1983). Such dissociation effects (i.e., knowledge that can be activated in one context remains inert in another) are exacerbated by time pressure (Cook and Woods, 1994). As interruptions accumulate and time pressure intensifies, the knowledge needed to address the situation at hand may not be activated, producing a situation such as that highlighted by Weick (1993b), in which retreating wildland firefighters failed to identify dropping their tools as an appropriate response to a rapidly advancing fire. Error rates also increase as less pertinent knowledge is accessed (Raufaste, Eyrolle, and Marine, 1998) and as increasingly poor attention management and goal selection direct attention to the wrong problems (Cook and Woods, 1994). Finally, the accumulation of interruptions pending, because it increases the number and type of issues to be resolved, precipitates conflicts among operational goals (e.g., safety and productivity, speed and accuracy) and creates double binds in which people may feel damned if they do and damned if they don't take a certain course of action (Cook and Woods, 1994). Thus, an accumulating stock of interruptions pending will eventually cause the rate at which interruptions are successfully resolved to collapse, because it simultaneously compromises attention management, knowledge activation, and the ability to trade off competing goals.

The links among the stock of interruptions pending, the Yerkes-Dodson curve, and the net resolution rate are shown in figure 2, which provides a pictorial overview of our model. To understand this diagram (and the underlying model), consider the situation faced by the KLM crew just prior to the Tenerife accident. As they attempted to execute the normal taxi and take-off procedure, they were interrupted by a constant stream of noisy and non-standard transmissions from

Figure 2. Model structure.



the tower. This stream of communication constitutes the interruption arrival rate. Such interruptions are not handled instantaneously but, instead, accumulate in the stock of interruptions pending until they are interpreted and then resolved. The stock of pending interruptions coupled with the time horizon over which they must be resolved determines the desired interruption resolution rate. Given their desire to leave within duty-time constraints, the KLM crew had an aggressive timetable for resolving these interruptions, leading to a short *desired resolution time* (1 minute). The comparison of the outstanding stock of unresolved communications and the resolution time then generates the *desired resolution rate,* the average pace of interruption resolution required to achieve the desired resolution time. As the number of outstanding interruptions rises, the desired resolution rate also rises. Formally,

$$\text{Desired resolution rate}(t) = \text{Interruptions pending}(t)/\text{Desired resolution time}.$$

The rationale for the increase in *desired resolution rate* is straightforward: The greater the number of non-standard directives from air traffic control, the greater the rate at which the crew must complete them to reduce the stock of pending interruptions.

We modeled *stress* as arising from a mismatch between the desired resolution rate and the rate at which interruptions are normally resolved. Formally,

$$\text{Stress}(t) = \text{Desired resolution rate}(t)/\text{Normal resolution rate}$$

(where normal resolution rate = 10 interruptions per minute). In the case of the KLM crew, responding to the increasing number of tower transmissions (along with the other non-standard features of the situation) meant they needed to work faster than normal (their desired resolution rate increased), and this need created stress (desired resolution rate > normal resolution rate). Finally, as the figure also highlights, the link between stress and the interruption resolution rate (performance) is captured by the Yerkes-Dodson curve. Initially, as stress increases, performance also increases (the system is operating on the upwardly sloping portion of the Yerkes-Dodson curve). Eventually, however, the peak of the curve is reached, and further increases in stress cause a decline in performance.[3]

Positive and Negative Effects of Stress: A Feedback Representation

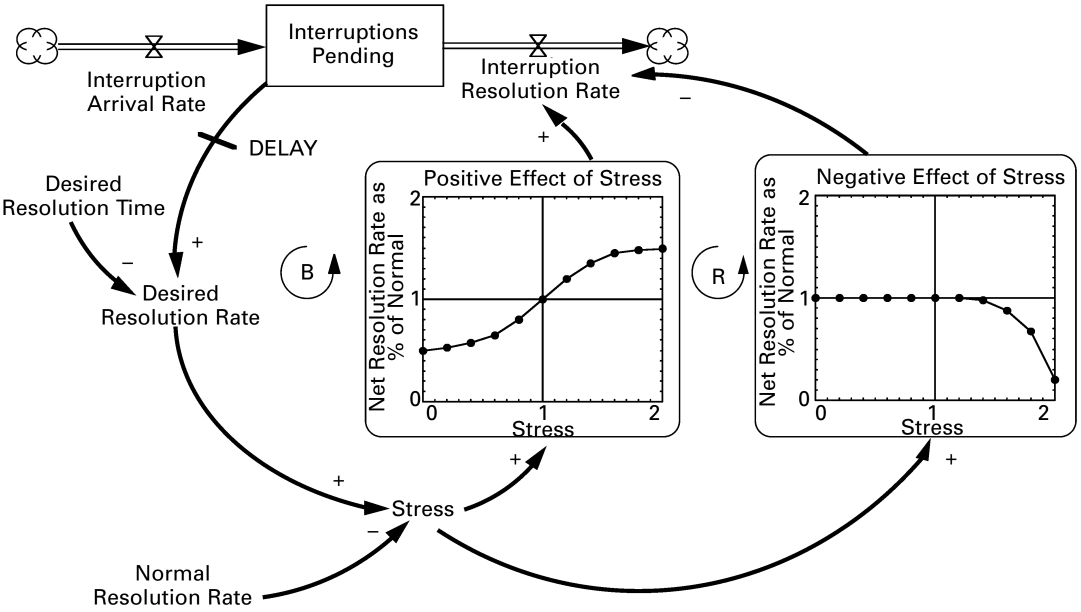
Although not immediately obvious, the dynamics of this system are quite different depending on whether it operates on the upward- or downward-sloping portion of the Yerkes-Dodson curve. To isolate these different dynamics, in figure 3 we separate the Yerkes-Dodson curve into its upward- and downward-sloping components. While the mathematical structure of the model is identical to that in the previous representation, this view highlights the two different feedback processes at work. To understand the upward-sloping portion, imagine the Combat Information Center on the *Vincennes.* The team must handle an increased stream of interruptions brought about by the nearby gun battles and their own malfunctioning forward gun. The accumulating stock of unresolved interruptions leads to an increase in the desired resolution rate and, therefore, stress. The increase in stress pushes the system farther up the Yerkes-Dodson curve, causing a positive change in the net resolution rate. The increased net resolution rate drains the stock of interruptions pending, thereby offsetting the initial increase. Thus, when operating in the upward sloping portion of the curve, the feedback process outlined above acts as a balancing or deviation counteracting feedback loop, represented by the “B” in the loop’s center: as the stock rises, stress grows, and performance increases, thereby draining the stock of outstanding interruptions and offsetting the initial increase.

If the stress level rises enough to push the system into the downward-sloping portion of the curve, however, the system’s response changes considerably. For example, the *Vincennes’* crew was already handling a significant volume of radio, radar, computer, and vocal inputs and queries when the appearance of an unidentified, possibly hostile aircraft caused a further increase in the stock of interruptions pend-

---

[3] Formally, we capture the linkage between stress and the resolution rate with the following equation: Net interruption resolution rate(t) = Normal resolution rate * Effect of stress on performance[Stress(t)], where the *effect of stress on performance*[•]is a nonlinear function, shown in figure 2, depicting the U-shaped relationship between stress and performance.

Figure 3. A feedback representation of the model's structure.



Note: Arrows indicate the direction of causality. Plus or minus signs at arrow heads indicate the polarity of relationships: a plus sign denotes that an increase in the independent variable causes the dependent variable to increase, ceteris paribus, and a decrease causes a decrease. Similarly, a minus sign indicates that an increase in the independent variable causes the dependent variable to decrease (see Sterman, 2000).

ing. As before, a rise in the stock of unresolved interruptions increases the desired resolution rate and creates more stress. But now, because the system is operating in the downward-sloping portion of the Yerkes-Dodson curve, the added stress causes a decline in the net resolution rate. A decline in the net resolution rate in this situation is akin to plugging the drain in a bathtub when the water is still running; it causes the level of unresolved interruptions (the water level) to grow further. Rather than offsetting the change in the number of unresolved interruptions, when operating in the downward-sloping section, the system amplifies it. Here, the dynamics of stress and performance do not perform a regulatory function but, instead, amplify changes in stress in a reinforcing feedback process (labeled with the loop identifier "R"). Thus, as the system moves from the upwardly sloping to the downwardly sloping portion of the Yerkes-Dodson curve, there is a shift in loop dominance. Whereas, when operating in the upward section, the balancing loop dominates the system's behavior, once in the downward-sloping section, the reinforcing loop determines its dynamics. And, unlike in linear dynamic systems, where the relative strength of various feedback loops is constant, in nonlinear systems like this one, the loop dominance can change. Such a shift is central to understanding the dynamics of quantity-induced disaster. To develop the main insights, along with their implications for understanding crises, we begin with a set of stylized simulation experiments that highlight the system's most important (and counterintuitive) dynamics.

Illustrating System Properties with Pulse Tests

In the experiments that follow, we assumed that the system begins in equilibrium. Formally, this means that the inflow to the stock of interruptions pending (the interruption arrival rate) equals the outflow (the net interruption resolution rate). Practically, equilibrium represents a system operating as designed in steady state. To understand its behavior in disequilibrium situations, we perturbed the system using a variety of test inputs (changes in the interruption arrival rate). We began with a stylized test input, a temporary, one-time increase in the interruption arrival rate. Such "pulse" tests are not particularly realistic but are widely used in the analysis of dynamic systems because they often provide a clear picture of how a system behaves in disequilibrium situations (Sterman, 2000). Figures 4a and 4b show the response of the system to two different pulse inputs. The first increases the interruption arrival rate by 100 percent for one minute; the second increases it by 120 percent for the same duration.

The system's response to the 100-percent increase in the arrival rate shown in figure 4a causes a sharp increase in the stock of interruptions pending. The growth in interruptions pending increases stress (not shown), which triggers a rise in the net resolution rate. Once the net resolution rate exceeds the arrival rate (meaning the outflow is greater than the inflow), the stock of interruptions pending begins to fall. The decline in the number of unresolved interruptions reduces stress and returns the net resolution rate to its steady state value. Due to the delay in perceiving changes in the stock of interruptions, the system actually undershoots slightly, briefly oscillating before returning to its initial equilibrium. Here the feedback between stress and performance plays a regulatory role, allowing the system to accommodate the temporary increase in the interruption arrival rate.

Contrast this outcome with the experiment shown in figure 4b, in which the workload is increased by 120 percent of the steady-state value. The larger pulse causes a greater increase in the stock of interruptions pending. As before, the growing stock of unresolved interruptions creates more stress and increases the net resolution rate. In this case, however, the increase is short lived. As interruptions continue to accumulate, stress builds, and the net resolution rate begins to fall. The fall in the net resolution rate results from the declining influence of the balancing loop and the growing strength of the reinforcing loop. Once the net resolution rate falls below the steady-state arrival rate (at approximately minute seven), the stock of interruptions pending shoots up, raising stress and causing the system to collapse.

The pulse experiments highlight three important features of the system's dynamics. First, the relationship between system performance and the interruption arrival rate is not well captured by the proportional logic embodied in statements like "the more there are, the worse things get." Nor does it follow the pattern suggested by the Yerkes-Dodson curve, a steady increase in performance followed by an equally paced collapse. Instead, despite the relative similarity of the two

Figure 4a. System response to a one-time increase in the interruption arrival rate of 100 percent.

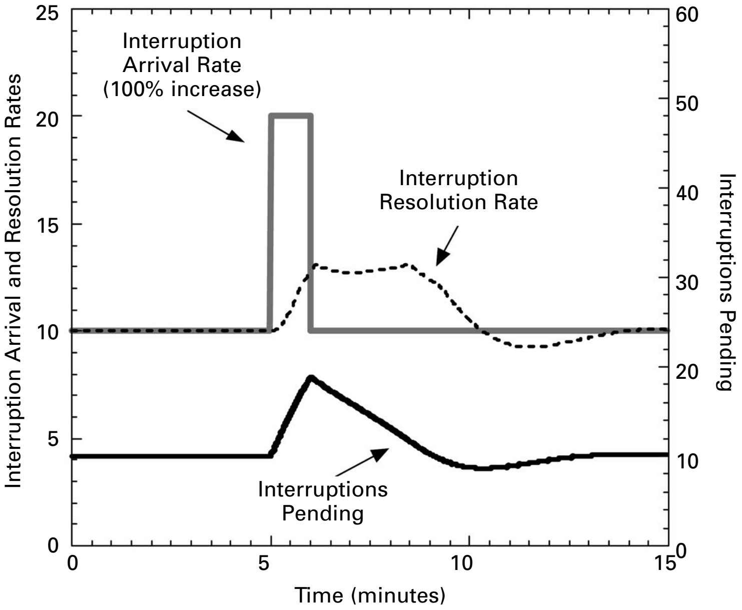
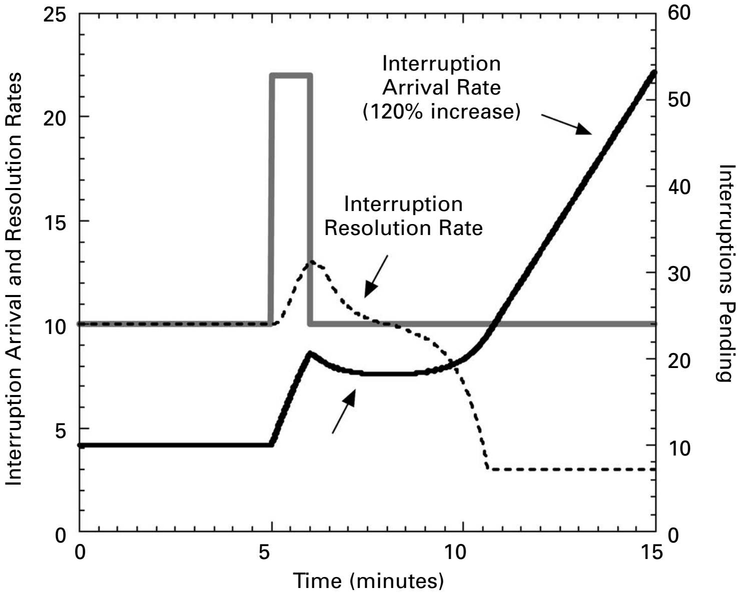


Figure 4b. System response to a one-time increase in the interruption arrival rate of 120 percent.



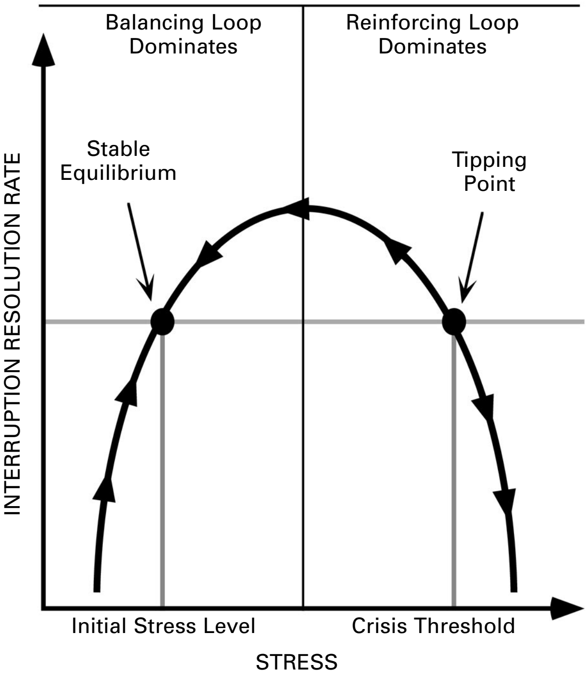
experiments, they generate qualitatively different dynamics: the system ably accommodates the 100-percent increase, easily returning to its pre-pulse equilibrium, but the 120-percent increase results in a rapid collapse. Second, as confirmed by a more comprehensive set of pulse simulations not shown here, an infinitesimally small change in the size of the pulse can mean the difference between survival and collapse (a graph of the results is available in the technical appendix). Thus, the experiments suggest that the relationship between interruptions and performance, rather than being proportional, is better captured by the notion of a critical threshold, which, once exceeded, causes a fundamental change in the system's behavior. Finally, the pulse experiments also show that this shift can be occasioned by a transient increase in the number of incoming interruptions. In all the experiments, the increased rate of interruption arrival lasts for only one minute. While it should come as little surprise that permanently overloading a system with interruptions eventually results in crisis, the experiment suggests that such systems can be more fragile; in the system we study, a temporary increase in workload causes a permanent decline in the system's performance.

Tipping Points and the Dynamics of Quantity-induced Crises

The pulse experiments suggest that systems facing an ongoing stream of non-novel interruptions have thresholds of accumulation beyond which their response to new demands fundamentally changes. Whereas when they operate below that threshold they are resilient, easily accommodating changes in the number of incoming interruptions, once the threshold is crossed, performance rapidly collapses. Figure 5 helps explain this divergent behavior.

Our characterization of the system's dynamics begins with the curvilinear relationship between stress and performance (where performance is represented as the net interruption resolution rate) embodied in the Yerkes-Dodson curve. Capturing the dynamics created by this relationship requires three additions. First, we highlight the shift in loop dominance that occurs when the system reaches the peak of the Yerkes-Dodson curve. Second, we show the possible equilibria, represented in figure 5 by solid black dots. An equilibrium exists whenever the steady-state interruption arrival rate pro-

Figure 5. The Yerkes-Dodson curve in dynamic environments.



duces enough stress to yield an equivalent net resolution rate (i.e., the inflow to the stock equals the outflow). The system has two such points, one on the upward-sloping segment and one on the downward-sloping portion. While both points represent equilibria, their dynamic characteristics (i.e., how the system behaves when it is perturbed from those equilibria) are quite different.

Third, to capture the differing dynamic characteristics of the two equilibria, we add arrows showing the direction or trajectory of the system in disequilibrium situations. In the equilibrium located in the upward-sloping portion of the Yerkes-Dodson curve, the trajectory arrows point toward the equilibrium point. Formally, this equilibrium is stable, meaning that small deviations from it are counteracted by the system's dynamics. The equilibrium's stability results from its location in a region where the balancing loop's stabilizing force dominates. The dynamics the system generates when operating near this equilibrium are highlighted in the first pulse experiment: when the stock of interruptions pending is increased, the resulting growth in stress increases the net resolution rate, thereby offsetting the larger number of unresolved interruptions and bringing the system back to equilibrium.

In contrast, the second equilibrium exists in a region where the reinforcing loop dominates the behavior of the system. Formally, this equilibrium is unstable, meaning that small deviations from it, rather than being counteracted, are amplified (note that the trajectory arrows point away from this equilibrium). Because it is unstable, it is extremely unlikely that the system will ever settle at this equilibrium. Nonetheless, it plays a critical role in determining the system's behavior because it is the threshold of stress at which the system undergoes the transformation highlighted in the pulse experiments, a transformation critical to understanding the role of quantity in precipitating crisis.

Although not widely acknowledged in the organizational literature, reinforcing (or positive) feedback processes can often work in one of two directions. In the system we model, the reinforcing loop created by the downward-sloping portion of the Yerkes-Dodson curve can act, depending on the level of stress, as either a virtuous cycle (fewer unresolved interruptions, less stress, and an increasing net resolution rate) that drives the system back to its initial equilibrium or as a vicious cycle (more unresolved interruptions, increasing stress, and a declining net resolution rate) that drives the system toward collapse. The unstable equilibrium represents the point at which the positive loop changes direction. Before the system reaches this threshold, the reinforcing loop enhances stability—the declining number of interruptions pending reduces the amount of stress and increases the net resolution rate—and, as the arrows indicate, drives the system back toward its initial equilibrium. Once the unstable equilibrium is crossed, however, the reinforcing loop, rather than pushing the system to stability, becomes an engine of disaster, driving the system into a crisis of escalating interruptions pending, intensifying stress, and declining performance.

The significance of this threshold, and consequent change in the system's behavior, can best be understood with an analogy to a concept arising from the study of epidemics. In models of infectious-disease propagation, the tipping point represents the size of the infected population required for a disease episode to become an epidemic (Murray, 1993). The tipping point is important because, once crossed, the dynamics of the system, and therefore the problem faced by public health officials, undergo a fundamental transformation. What was once just a particular disease episode that would have quickly run its course suddenly becomes an epidemic threatening to infect the entire susceptible population. Similarly, in our model, the unstable equilibrium represents a kind of tipping point. When operating below the unstable equilibrium, the natural regulatory effects of stress dominate, and participants are likely to feel that they are in control, easily accommodating changes in the arrival rate with adjustments in productivity. When the tipping point is crossed, however, the system's response to additional interruptions fundamentally changes. Once beyond the tipping threshold, the system that initially seemed resilient and amenable to human intervention becomes pre-programmed for crisis. What once felt like an orderly, rational world suddenly becomes a system seemingly beyond human control.

Variability and Performance

Having developed the central notion of our analysis—that organizations can have tipping points beyond which their response to interruptions changes dramatically—we turn to how the resulting dynamics play out in real-world situations. The rate at which normal routines and procedures are interrupted can be quite variable. Just as air traffic controllers face significant and ongoing variability in the number of planes they manage throughout the day, questions in an academic presentation do not arrive at perfectly spaced intervals. To capture these real-world demands more realistically, we subjected the system to ongoing variation, or noise, in the interruption arrival rate.

To highlight how such variability affects system performance, we begin with two simulations in which all parameters are identical except that, in the second simulation, we double the variation in the rate of interruption arrivals (the mean arrival rate does not change). Figure 6a shows the first experiment. The system easily accommodates the variation in the number of interruptions. The arrival rate does rise significantly at approximately minute sixty, but, because the tipping point is not crossed, the system does not descend into crisis. In the second experiment, shown in figure 6b, despite having the same average number of arrivals, doubling the deviation in the arrival rate poses a significant problem. The larger swings in the arrival rate push the system farther away from its initial equilibrium in both directions. The downswings leading to reductions in stress and downward departures from the stable equilibrium have little effect on system behavior. The balancing loop continues to regulate the system, constantly driving it back to its starting point. But the excursions on the right side of the equilibrium are more problematic. The increased arrival rate at approximately minute sixty pushes

Figure 6a. System response to variation in the interruption arrival rate.

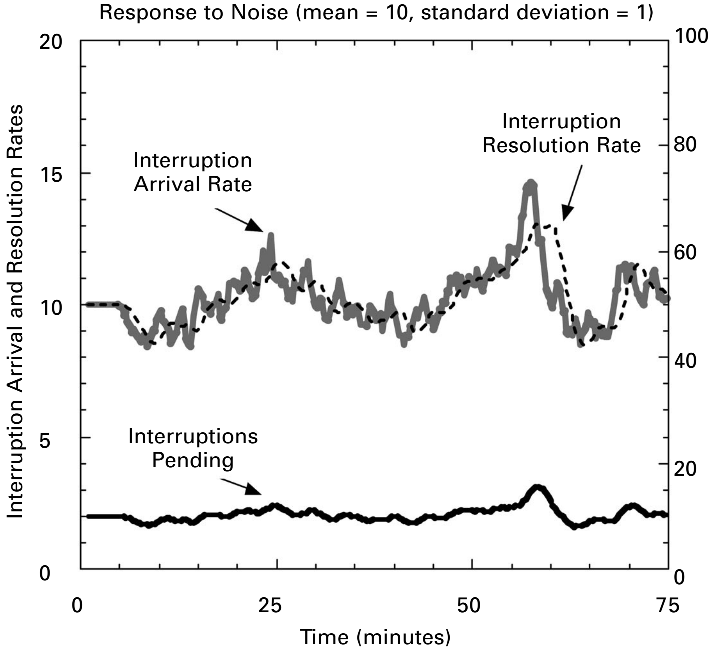
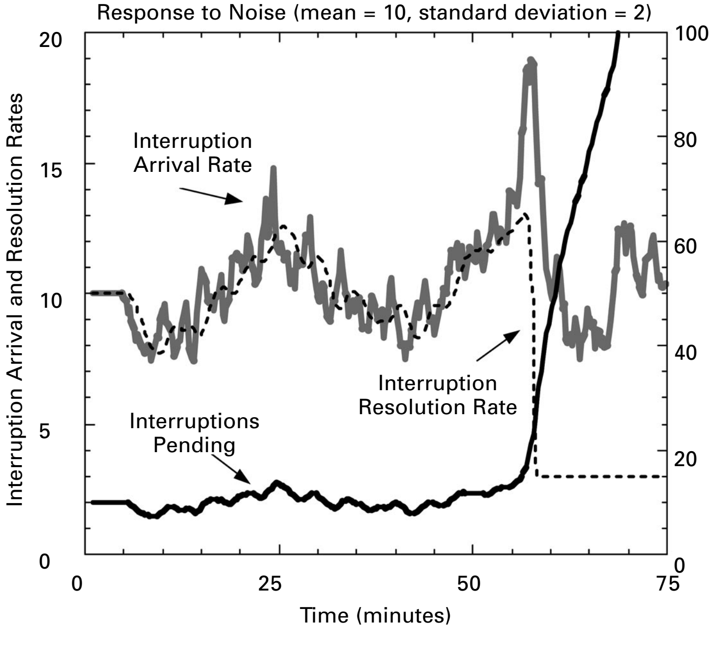


Figure 6b. System response to greater variation in the interruption arrival rate.

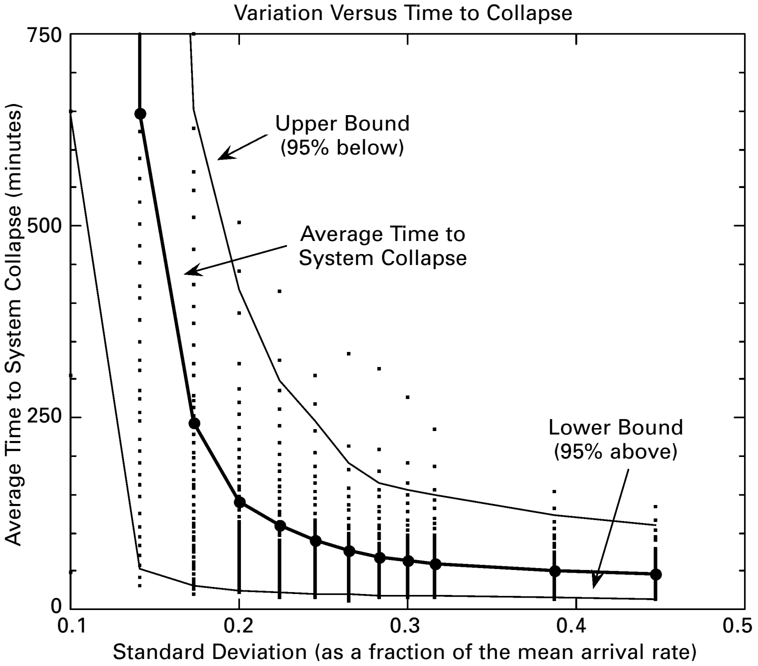


the system over the peak of the Yerkes-Dodson curve and then past the tipping point. Once that threshold is crossed, the net resolution rate falls below the arrival rate, and the stock of interruptions pending explodes in a vicious cycle of increasing stress and declining performance.

The experiments highlight two important features of the system's dynamics. First, the existence of a tipping point means that the variability of the arrival rate is a significant determinant of a system's susceptibility to crisis. The dynamics of the system are such that the tipping point only needs to be exceeded for a moment to create disaster. The wider the variation in the arrival rate, the more likely the tipping threshold will be crossed. Though the average number of arriving interruptions is the same in both experiments, the second experiment ends in disaster, while the first does not. Second, performance collapses very rapidly once the tipping threshold is exceeded. In the second simulation, although it took more than fifty minutes to reach the tipping point, once crossed, the self-reinforcing cycle of increasing stress and declining performance drives the system into collapse in just a few minutes.

Because these experiments show only two sequences of interruption arrivals, we used extensive Monte-Carlo analysis to confirm the relationship between variability and susceptibility to crisis.[4] We performed 1,000 simulations for each of twelve different standard deviations of the arrival rate process (ranging from 10 to 50 percent of the mean arrival rate). To summarize this wealth of data, we calculated and plotted the mean time to crisis and upper and lower confidence bounds for each of the selected standard deviations. Figure 7 depicts the relationship between the variability in the arrival rate and the average time required for the system to descend into crisis. As the variance in the arrival rate increases, the average survival time of the simulated systems falls exponentially. When the variance in the arrival rate is low, the variation in survival time is extremely large. When the arrival rate variance was lowest, survival times ranged between 78 and 15,000 minutes. This implies that even systems in low-variability environments face the possibility of a quantity-induced crisis; a closely grouped sequence of small interrup-

Figure 7. Monte-Carlo analysis of variability in the interruption arrival rate versus system survival time.



---

[4] Details of the Monte-Carlo analysis are available from the second author.

tions (a run of bad luck), no matter how unlikely, can push the system over its tipping point. For example, extensive training in simulators has provided pilots with the skills necessary to accommodate a wide range of unlikely but potentially catastrophic events, including the complete loss of normal control (Helmreich and Foushee, 1993). Yet, as the Tenerife case highlights, despite this training, a tight grouping of mundane interruptions to normal procedures can push the system into a pathological regime that degrades performance both so significantly and so quickly that a disaster that seems impossible under normal circumstances becomes possible.

Context and Susceptibility to Crisis

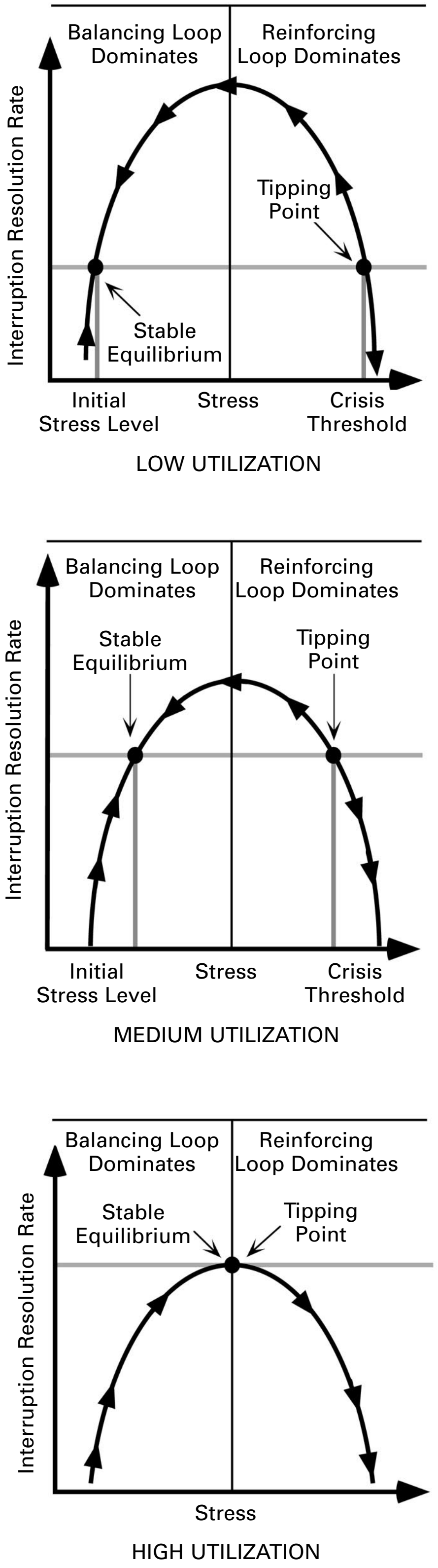
The existence of a tipping point and the consequent vulnerability to variation in the arrival rate raises the question of how many interruptions are required to push the system into crisis. As highlighted in figure 8, the balance between handling capacity (the normal net interruption resolution rate) and the steady-state arrival rate determines the number of interruptions required to tip the system into crisis. When the average arrival rate is low relative to handling capacity, the system operates at a low point on the Yerkes-Dodson curve. Consequently, a very large shock is required to push the system over its tipping point. In this situation, the system can handle substantial variability without descending into crisis. As the steady-state arrival rate increases, or the resolution capacity declines, however, both the stable equilibrium and the tipping point climb the Yerkes-Dodson curve, causing the distance between them to shrink. When the system reaches its maximum output, the initial equilibrium and the tipping point converge to a single point, at which even the slightest perturbation from that equilibrium causes a downward spiral into crisis. Thus, the system's ability to accommodate changes in the arrival rate without descending into crisis is a function of the balance between incoming interruptions and the system's capacity to handle them.

Building on this insight, we operationalize "context" in our model as the average number of incoming interruptions in a given environment. For example, in the case of the *Vincennes,* the ongoing gun battles, malfunctioning equipment, and other interruptions are likely to have pushed the system close to its tipping point. In such a context, even the slightest additional interruption, easily handled under other circumstances, was probably sufficient to push the system over the tipping point. If that threshold was crossed, information processing capability would have declined so rapidly that mistaking a commercial airliner for an F-14 became a distinct possibility (cf. Roberts and Dotterway, 1995).

Individual, Group, and Organizational Responses

While the general dynamics we capture are likely to be present at multiple levels of analysis, there are important differences in how individuals, groups, and organizations might respond to an accumulating stock of non-novel interruptions. At the individual level, the experimental literature describes a number of cognitive responses to overload. Such adjustment processes include omission, simply ignoring some signals;

Figure 8. System structure under three different levels of resource utilization.



error, handling them incorrectly; queuing, delaying some interruptions in favor of others; filtering, attending to important interruptions first; abstracting, processing interruptions by glossing over details; escape, reducing the rate of interruption arrival; and chunking, resolving groups of interruptions rather than focusing on them in sequence (Miller, 1978: 146). Three of these, omission, error, and queuing, are already captured in our analysis. In the discussion, we address escape, which, if feasible, is an effective strategy for assuring the survival of the system. The final three, filtering, chunking, and abstracting, which increase effective resolution capacity, play similar roles in the dynamics discussed so far.

By improving the net resolution rate, all of these responses increase the distance between the stable equilibrium and the tipping point. If filtering, chunking, and abstracting can be executed effectively, they make the system more robust. But these adaptations, while changing its location, do not eliminate the tipping point. The tipping point arises from two features of the system we model: (1) unresolved interruptions do not disappear but, instead, accumulate; and (2) the accumulation eventually causes a drop in performance. As long as mounting stress eventually degrades performance, the system has a tipping point. There is no evidence to suggest that individual-level responses prevent the inevitable slide down the right side of the Yerkes-Dodson curve caused by ever-increasing levels of stress. Thus, while strategies such as chunking and abstracting increase a system's resilience, the susceptibility to a quantity-induced crisis remains.

At the group level, handling capacity, coordination, and control also affect the location of the tipping point but, again, are unlikely to eliminate it. Adding people to the group can increase the distance between the stable equilibrium and the tipping point. If the level of interruptions continues to increase, however, people will eventually become overwhelmed and crisis will ensue. Further, interruptions are rarely independent events; resolving them often requires coordination among group members, thereby adding additional steps in the resolution process. Handling capacity is not likely to scale directly with the addition of people, though improved group dynamics may also increase handling capacity. For example, with high levels of synergy, groups can optimize capacity and coordination (Hackman, 1989). Similarly, the constriction of control and information processing that accompany high levels of stress, which can lead to reactions like groupthink (Janis, 1982), are beneficial when dealing with non-novel interruptions (Staw, Sandelands, and Dutton, 1981). Whether the result of synergy or groupthink, group cohesiveness and uniformity of response are likely to increase the system's ability to handle non-novel interruptions.

Organizational-level responses can also shift the tipping point but, even under the most favorable conditions, are unlikely to eliminate it. Under increasing demands from the environment, organizations tend to centralize control and, like individuals, filter information more heavily (Staw, Sandelands, and Dutton, 1981). Filtering has the positive effect of reducing message overload and stress, allowing organizations to act.

Similarly, centralizing authority, and formalizing and standardizing processes and procedures (Staw, Sandelands, and Dutton, 1981) can speed the resolution of routine interruptions. None of these responses, however, prevents the eventual decline in the handling rate caused by an ever-mounting stock of unresolved interruptions. Thus, while our model does not answer the question of how the location of the tipping point changes at different levels of analysis, our results suggest that moving from the individual to the group to the organizational level does not change the features required for its existence.

DISCUSSION

Our analysis offers two contributions to understanding the resilience of organizational systems. First, we highlight quantity as a basis for disaster. While a substantial body of work has examined how novel events can precipitate disaster, we show how such catastrophic outcomes can be the result of an overaccumulation of mundane events, any of which, on its own, poses little threat to the organization. Second, we show that the relationship between the quantity of interruptions and the organization's ability to handle them is far more complex than casual intuition might suggest. Organizations facing an ongoing stream of routine but survival-threatening interruptions have tipping points, thresholds of accumulated interruptions beyond which performance rapidly collapses. These organizations can survive for extended periods of time, ably accommodating the interruptions they face, until a closely grouped sequence (a run of bad luck) pushes the system over its tipping point, rapidly degrading performance to the point that disaster is almost inevitable.

The model from which these insights are derived is far simpler than any real-world system. We have not modeled the task being interrupted, nor have we captured the dynamics specific to different levels of analysis. More importantly, in the interest of studying quantity, we have not captured the role of novelty. Despite these limitations, the model has two important implications for both the prevention of crises and future research on them: (1) understanding disaster and its precursors requires considering both the novelty and the quantity of interruptions, and (2) it is a grave error to assume that both types of events occasion similar dynamics. To develop these implications, we offer two constructs for thinking about disaster and resilience: a novelty-induced crisis and a quantity-induced crisis.

A novelty-induced crisis results from an interruption for which an organization does not have an appropriate response within its repertoire. A crisis arises in this situation solely due to the interruption's incomprehensibility, not from the time frame over which it must be resolved. Autonomic arousal and the dynamics of stress we have discussed play little role in this type of crisis. Examples include chronic diseases for which there are no known cures and social problems like homelessness whose negative consequences, while lethal, play out over long time scales. In contrast, a quantity-induced crisis is created by a series of interruptions, each of which can be resolved using a known response (although time is required to identify and execute that response), that overwhelms information processing capacity and creates a vicious cycle of increasing stress and declining performance. Examples of this type range from the famous "Candy Factory" episode of the television show *I Love Lucy,* in which Lucy is overwhelmed by the increasing speed of the chocolate factory assembly line, to the accident at Tenerife.[5]

The distinction between the two types of crises has important implications for how organizations attempt to prevent them and what they do once one is in progress. When organizations are confronted with novelty, the proposed solution often lies in enlarging the repertoire of organizational responses, building resilience and the ability to cope with "surprises in the moment" (Weick, Sutcliffe, and Obstfeld, 1999: 100). Organizational theorists often argue that people must step back from the situation at hand, revisit their core assumptions, reframe the situation, recombine existing procedures and routines into alternative responses (e.g., improvisation), and engage in some type of higher-order evaluation, such as double-loop learning (Argyris and Schön, 1974). Such actions make sense when the source of difficulty is novelty, but they are problematic when the quantity of interruptions is an issue.

As examples like Tenerife demonstrate, when a crisis is quantity-induced, people often do not recognize an impending disaster until it is too late. The problem arises from the nonlinearity of the system and the rapidity of collapse once a crisis is initiated. When operating below the tipping threshold, the system provides powerful evidence to participants that, with just a little more effort, they can handle just a few more interruptions. After all, until the threshold is reached, this strategy works perfectly. With time, people become increasingly confident in their ability to offset variations in the interruption arrival rate with changes in their productivity. The system teaches them that they are in control, and they are—until the system crosses the tipping point and they are bushwhacked by a vicious cycle of declining performance and accumulating interruptions. These difficulties are further exacerbated by the fact that as the system approaches its tipping point, the ability to sense impending disaster is likely to decline. The accumulating stock of interruptions, coupled with mounting autonomic arousal leads to perceptual narrowing, less activation of relevant knowledge, worsening attention management, and poor strategic choices, all of which limit the ability both to realize a crisis is at hand and to handle it appropriately.

Even more problematic, when people recognize an impending crisis, attempts to implement an alternative response can often make the situation worse rather than better. The stepping back recommended in the literature on novelty-based crisis takes time, causing the net resolution rate to decline temporarily while people undertake these activities. This implies that the responses themselves can push the system over the tipping point. Our analysis shows that once interruptions start to accumulate above their normal level, reframing the situation, if it takes time and temporarily reduces the net resolution rate, only increases the rate of accumulation and

---

[5] This episode originally aired on 15 September 1952. A video clip is available at http://www.lucylibrary.com/Pages/ill-guide-2.html.

hastens the system's collapse. Once the system is stuck in a vicious cycle of accumulating interruptions, increasing stress, and declining performance, it's too late for deliberative double-loop learning.

Ironically, as a consequence of these two limitations and in sharp contrast to the recommendations of those who have studied novelty-induced crises, unquestioned adherence to preexisting routines may be the best way to prevent the overaccumulation of pending interruptions. For example, climbing teams tackling major mountains such as Everest frequently impose a turnaround time rule: on the day a team attempts the summit, all members must turn around by a prespecified time, regardless of whether they have achieved their goal. Information processing and decision-making capability are severely restricted by low oxygen at extreme altitudes, so even a few interruptions, such as unplanned delays, can create dire problems. Experience suggests that, in such situations, it is better not to leave the turnaround time up to on-the-spot decision making. Accounts of a recent disaster on Everest highlight how violating such rules resulted in the death of nine people, including two experienced guides (Krakauer, 1997; Boukreev and De Walt, 1997).

The utility of unquestioned adherence to preexisting rules creates something of a paradox for preventing quantity-induced crises: rules like the turnaround time, which, to be effective, must be followed without question, are themselves the product of reflection and reframing. The resolution lies in recognizing that a strategy of questioning existing procedures that works so well when there is slack in the system can be disastrous when there is none. When climbing a formidable mountain like Everest, a climbing team's ability both to recognize and to overcome disruptions to its plans declines as it moves up the mountain. Thus, while it may be appropriate, even desirable, to rethink the turnaround-time rule while in base camp or when reflecting on lessons learned from the last trip, it should be followed without question on the day that climbers attempt the summit. Similarly, in the system we study, as the level of unresolved interruptions accumulates and the system operates at higher points on the Yerkes-Dodson curve, its resilience to additional interruptions declines. With few interruptions pending, the system is resilient to the costs of learning, reframing, and improvisation, but as the stock grows, eventually even the slightest perturbation can push the system over its tipping point.

The temporal interplay between double-loop approaches and unquestioned adherence to existing procedures is highlighted by Weick and Roberts' (1993) discussion of one evening's operations on an aircraft carrier. This case also demonstrates the value of Miller's (1978) escape response, reducing the outstanding number of accumulated interruptions. On the night in question, an unusually large number of unanticipated mechanical problems occurred in aircraft waiting to land. At one point, there were several planes waiting to take off and five planes in the final approach pattern waiting to land, three of which were experiencing mechanical difficulty. As demands on the system increased, the aircraft waiting to land with the most severe difficulties was eventually instructed to reroute to a mainland airstrip, but the pilots had to eject when it ran out of fuel. The plane was lost, but both pilots were unhurt. Weick and Roberts reported this as an example of a system failure caused by the breakdown of "heedful interrelating," communicating about and rapidly adjusting to changing conditions.

While the actions prior to sending the plane to an alternate airstrip certainly constitute a breakdown in system performance, the actual decision to reroute the aircraft may have been instrumental in defusing a disaster already in progress. As Weick and Roberts (1993: 373) wrote, "There is a limit to heedfulness . . . and on that night this ship was at that limit. The system was overloaded and the situation was one that managers of high-technology weapon systems worry about all the time. They call it OBE (overcome by events)." Given this characterization, the system may have been approaching its tipping point. Sending one plane to an alternative landing constituted a reduction in the number of outstanding interruptions faced by the system, perhaps moving it back into a more resilient regime; the other aircraft did land without incident. And, while the strategy cost at least $38 million (the price of the lost airplane), it may have kept the system from descending further into crisis, preventing higher financial costs and loss of life. Thus, while Weick and Roberts' analysis highlights the value of allowing people the latitude to adapt to the changing requirements created by a complex system like an aircraft carrier, our analysis raises an important caveat: fixed rules and procedures can prevent the system from entering a regime in which the participants' ability to execute the higher-order thought processes required for such "heedful" interaction is severely restricted. The quantity-induced-crisis construct implies that organizations prone to such events face the difficult task of concurrently cultivating the ability to challenge existing procedures, follow them without question, and determine the conditions under which each action is appropriate.

The complexity of preventing and managing crises increases further when one realizes that many (if not most) crises combine elements of both quantity and novelty. In many cases, organizations must cope with interruptions that are both novel and must be resolved quickly. Weick's (1993b) analysis of the fire at Mann Gulch provides one example: wildland firefighters were confronted not only by novelty (a fire far larger than they expected) but also by the demands of quantity; the fire was approaching rapidly, dictating that a response be implemented quickly. Organizations facing such mixed crises must cultivate a complex set of skills, capabilities that have received little attention in the existing literature.

Our analysis suggests at least two directions for future research on enhancing organizational resilience. First, while it is often fairly obvious when an organization is confronted by novelty, quantity-induced crises are harder to recognize. Until the tipping point is reached, things look fine. Future research could be profitably focused on helping organizations assess their capacity to handle interruptions and develop systems to signal when the level of unresolved interruptions reaches a critical level. Our analysis provides one hint as to how such a system might be developed. Specifically, quantity-induced crises are preceded by a predictable pattern of events. Initially, the net resolution rate increases in response to mounting interruptions and stress. The improvement in performance continues until the system reaches the peak of the Yerkes-Dodson curve. If stress continues to increase, however, the system moves over the peak, and performance (the net resolution rate) falls, despite the increasing level of interruptions pending. Thus, our analysis suggests that quantity-induced crises may have an identifiable "temporal signature": a rising level of unresolved interruptions and an initial improvement in performance, followed by a decline. If this pattern withstands empirical scrutiny, it may provide the basis for an early-warning system that allows managers to defuse what might otherwise be a crisis already in progress.

A second avenue for future research relates to the challenges facing organizations that must trade off responding to quantity and novelty: novelty requires time-consuming reframing and invention, while quantity requires adherence to rules that prevent the overaccumulation of outstanding interruptions. Research on effective management of simulated medical crises (Rudolph, 2002) provides one example. A routine and central component of medical resident training is learning to generate a differential diagnosis, a list of diagnoses about what could be wrong with a patient. Medical residents new to crisis management, however, frequently abandon this process, making erroneous therapeutic decisions based on both a truncated review of symptoms and a truncated differential diagnosis. Although this is understandable given the time pressure of the situation, preliminary observations suggest that requiring trainees to think aloud and solicit input on forming a list of four or five symptoms and hypotheses helps them activate inert knowledge and more effectively shift attention between details and the big picture, often pointing them to a better solution within less than two minutes. This approach forces an explicit balance between the competing challenges of resolving the problem quickly (a quantity-related concern) and correctly (a novelty-related concern) and makes questioning existing frames a standard operating procedure.

Unfortunately, our analysis also suggests that the ability to cultivate both sets of skills (double-loop learning and unquestioned adherence to preexisting routines) and recognize the conditions under which each is appropriate is unlikely to improve with experience in non-crisis situations. The lessons people learn before they cross the tipping point are likely to misguide them once a crisis is underway. Experience uninformed by a thorough understanding of how this nonlinear system behaves under different conditions is unlikely to prepare people for the increased pace, discontinuous changes, and tighter coupling that characterize a system beyond its tipping point. Thus, our framework suggests that future research in organization studies can profit from a focus on understanding how groups and individuals maintain both skill sets along with the ability to determine which is most appropriate in a given situation.

REFERENCES

Argyris, C., and D. A. Schön
1974 Theory in Practice: Increasing Professional Effectiveness. London: Jossey-Bass.

Barry, J., and R. Charles
1992 "Sea of lies." Newsweek, July 13.

Bauemler, G.
1995 "On the validity of the Yerkes-Dodson law." Studia-Psychologica, 36: 205–209.

Boukreev, A., and G. W. De Walt
1997 The Climb. New York: St. Martin's Press.

Carroll, J. S.
1995 "Incident reviews in high-hazard industries: Sense making and learning under ambiguity and accountability." Industrial and Environmental Crisis Quarterly, 9: 175–197.

Coles, M. G.
1974 "Physiological activity and detection: The effects of attentional requirements and the prediction of performance." Biological Psychology, 2: 113–125.

Collyer, S. C., and G. S. Malecki
1998 "Tactical decision making under stress: History and overview." In J. A. Cannon-Bowers and E. Salas (eds.), Making Decisions under Stress: Implications for Individual and Team Training: 3–15. Washington, DC: American Psychological Association.

Cook, R. I., and D. D. Woods
1994 "Operating at the sharp end: The complexity of human error." In B. S. Bogner (ed.), Human Error in Medicine: 255–310. Hillsdale, NJ: Lawrence Erlbaum.

De Keyser, V., and D. D. Woods
1990 "Fixation errors: Failures to revise situation assessment in dynamic and risky systems." In G. Colombo and A. Saiz de Bustamante (eds.), Systems Reliability Assessment: 231–251. Amsterdam: Kluwer.

Fisher, S.
1986 Stress and Strategy. London: Lawrence Erlbaum.

Forrester, J. W.
1961 Industrial Dynamics. Portland, OR: Productivity Press.

Gentner, D., and A. L. Stevens
1983 Mental Models. Hillsdale, NJ: Lawrence Erlbaum.

Glaser, B. G., and A. L. Strauss
1967 The Discovery of Grounded Theory: Strategies for Qualitative Research. New York: Aldine.

Hackman, J. R.
1989 Groups That Work (and Those That Don't). San Francisco: Jossey-Bass.

Helmreich, R. L., and H. C. Foushee
1993 "Why crew resource management? The history and status of human factors training programs in aviation." In E. Wiener, B. Kanki, and R. Helmreich (eds.), Cockpit Resource Management: 2–45. New York: Academic Press.

Janis, I. L.
1982 Groupthink. Boston: Houghton Mifflin.

Klein, G.
1998 Sources of Power. Cambridge, MA: MIT Press.

Krakauer, J.
1997 Into Thin Air: A Personal Account of the Mount Everest Disaster. New York: Villard.

Louis, M. R., and R. I. Sutton
1991 "Switching cognitive gears: From habits of mind to active thinking." Human Relations, 44: 55–76.

Mandler, G.
1982 "Stress and thought processes." In S. Goldberger and S. Breznitz (eds.), Handbook of Stress: 88–164. New York: Free Press.
1984 Mind and Body. New York: W. W. Norton.

Miller, J. G.
1978 Living Systems. New York: McGraw-Hill.

Murray, J. D.
1993 Mathematical Biology. New York: Springer-Verlag.

Perrow, C.
1984 Normal Accidents: Living with High Risk Systems. New York: Basic Books.
1994 "Accidents in high-risk systems." Technology Studies, 1: 1–19.

Raufaste, E., H. Eyrolle, and C. Marine
1998 "Pertinence generation in radiological diagnosis: Spreading activation and the nature of expertise." Cognitive Science, 22: 517–546.

Reason, J.
1997 Managing the Risks of Organizational Accidents. Aldershot, UK: Ashgate.

Revelle, W., and D. A. Loftus
1990 "Individual differences and arousal: Implications for the study of mood and memory." Cognition and Emotion, 4: 209–237.

Roberts, K. H.
1990 "Some characteristics of one type of high reliability organization." Organization Science, 1: 160–176.

Roberts, N. C., and K. A. Dotterway
1995 "The *Vincennes* incident: Another player on the stage?" Defense Analysis, 11: 31–45.

Rudolph, J. W.
2002 "Fixation or hypothesis revision? Understanding and reducing impediments to situation awareness in emergent medical crises." Working paper, Carroll School of Management, Boston College.

Sagan, S.
1993 The Limits of Safety: Organizations, Accidents and Nuclear Weapons. Princeton, NJ: Princeton University Press.

Schulman, P. R.
1993 "The analysis of high-reliability organizations: A comparative framework." In K. H. Roberts (ed.), New Challenges to Understanding Organizations: 33–53. New York: Macmillan.

Shrivastava, P.
1987 Bhopal: Anatomy of a Crisis. Cambridge, MA: Ballinger.

Shrivastava, P., I. Mitroff, D. Miller, and A. Miglani
1988 "Understanding industrial crises." Journal of Management Studies, 25: 285–303.

Staw, B. M., L. E. Sandelands, and J. E. Dutton
1981 "Threat-rigidity effects in organizational behavior: A multilevel analysis." Administrative Science Quarterly, 26: 501–524.

Sterman, J. D.
2000 Business Dynamics. Chicago: Irwin-McGraw-Hill.

Strauss, A., and J. Corbin
1994 "Grounded theory methodology: An overview." In N. K. Denzin and Y. S. Lincoln (eds.), Handbook of Qualitative Research: 273–285. Thousand Oaks, CA: Sage.

Sullivan, S., and R. S. Bhagat
1992 "Organizational stress, job satisfaction and job performance: Where do we go from here?" Journal of Management, 18: 353–374.

Teigen, K. H.
1994 "Yerkes-Dodson: A law for all seasons." Theory and Psychology, 4: 525–547.

Turner, B. A.
1976 "The organizational and interorganizational development of disasters." Administrative Science Quarterly, 21: 378–397.
1978 Man-Made Disasters. London: Wykeham.

U.S. Senate Committee on Armed Services
1988 Investigation into the Downing of an Iranian Airliner by the USS *Vincennes*. Washington, DC: U.S. Government Printing Office.

Vaughan, D.
1996 The *Challenger* Launch Decision: Risky Technology, Culture, and Deviance at NASA. Chicago: University of Chicago Press.
1999 "The dark side of organizations: Mistake, misconduct, and disaster." Annual Review of Sociology, 25: 271–305.

Weick, K. E.
1979 The Social Psychology of Organizing. New York: McGraw-Hill.
1993a "The vulnerable system: An analysis of the Tenerife air disaster." In K. H. Roberts (ed.), New Challenges to Understanding Organizations: 173–198. New York: Macmillan.
1993b "The collapse of sensemaking in organizations: The Mann Gulch disaster." Administrative Science Quarterly, 38: 628–652.

Weick, K. E., and K. H. Roberts
1993 "Collective mind in organizations: Heedful interrelating on flight decks." Administrative Science Quarterly, 38: 356–381.

Weick, K. E., K. M. Sutcliffe, and D. Obstfeld
1999 "Organizing for high reliability: Processes of collective mindfulness." In R. I. Sutton and B. M. Staw (eds.), Research in Organizational Behavior, 21: 81–123. Stamford, CT: JAI Press.

Yerkes, R. M., and J. D. Dodson
1908 "The relation of strength of stimulus to rapidity of habit formation." Journal of Comparative Neurological Psychology, 18: 459–482.